PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/11, C07K 14/47, A61K 38/00

A2

(11) International Publication Number: WO 98/27205

(43) International Publication Date: 25 June 1998 (25.06.98)

(21) International Application Number: PCT/US97/23330

(22) International Filing Date: 17 December 1997 (17.12.97)

(30) Priority Data:

08/769,192	18 December 1996 (18.12.96)	US
08/783,401	13 January 1997 (13.01.97)	US
08/991,872	16 December 1997 (16.12.97)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

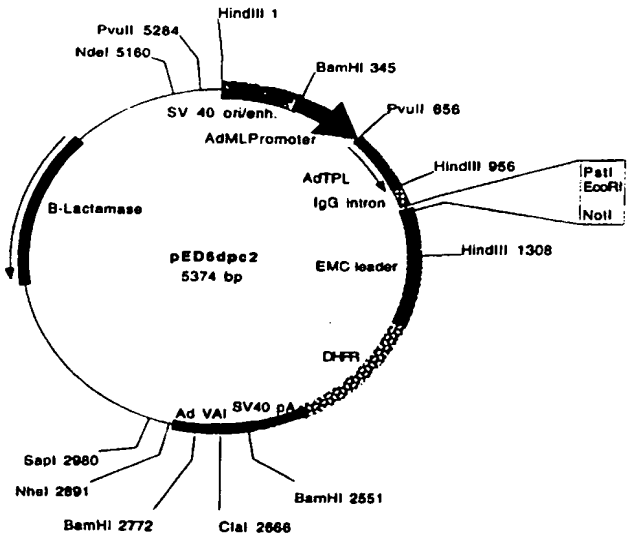
(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of the following applications: Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application 08/769,192), filed December 18, 1996; and Ser. No. 08/783,401, filed January 13, 1997; all of which are incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:2;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:2 from nucleotide 41 to nucleotide 760;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CB107_1 deposited under accession number ATCC 98279;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CB107_1 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:3;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:3 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:2 from nucleotide 41 to nucleotide 760; the nucleotide sequence of the full-length protein coding sequence of clone CB107_1 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CB107_1 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279. In yet other preferred

embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:3 from amino acid 127 to amino acid 240.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:2, SEQ ID NO:1 or SEQ ID NO:4 .

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:3;

(b) the amino acid sequence of SEQ ID NO:3 from amino acid 127 to amino acid 240;

(c) fragments of the amino acid sequence of SEQ ID NO:3; and

(d) the amino acid sequence encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:3 or the amino acid sequence of SEQ ID NO:3 from amino acid 127 to amino acid 240.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 374 to nucleotide 1108;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 500 to nucleotide 1108;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 387;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CG300_3 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CG300_3 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 374 to nucleotide 1108; the nucleotide sequence of SEQ ID NO:5 from nucleotide 500 to nucleotide 1108; the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 387; the nucleotide sequence of the full-length protein coding sequence of clone CG300_3 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CG300_3 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 23 to amino acid 57.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 23 to amino acid 57;

(c) fragments of the amino acid sequence of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 23 to amino acid 57.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 126 to nucleotide 3053;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 180 to nucleotide 3053;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 49 to nucleotide 382;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CJ145_1 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CJ145_1 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 126 to nucleotide 3053; the nucleotide sequence of SEQ ID NO:7 from nucleotide 180 to nucleotide 3053; the nucleotide sequence of SEQ ID NO:7 from nucleotide 49 to nucleotide 382; the nucleotide sequence of the full-length protein coding sequence of clone CJ145_1 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CJ145_1 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 87.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 87;

(c) fragments of the amino acid sequence of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 87.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 40 to nucleotide 342;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 127 to nucleotide 342;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 11 to nucleotide 181;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CJ160_11 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CJ160_11 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CJ160_11 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CJ160_11 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 40 to nucleotide 342; the nucleotide sequence of SEQ ID NO:9 from nucleotide 127 to nucleotide 342; the nucleotide sequence of SEQ ID NO:9 from nucleotide 11 to nucleotide 181; the nucleotide sequence of the full-length protein coding sequence of clone CJ160_11 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CJ160_11 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of

clone CJ160_11 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 7 to amino acid 48.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 7 to amino acid 48;

(c) fragments of the amino acid sequence of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone CJ160_11 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 7 to amino acid 48.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 180 to nucleotide 467;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 267 to nucleotide 467;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO20_1 deposited under accession number ATCC 98279;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO20_1 deposited under accession number ATCC 98279;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 180 to nucleotide 467; the nucleotide sequence of SEQ ID NO:11 from nucleotide 267 to nucleotide 467; the nucleotide sequence of the full-length protein coding sequence of clone CO20_1 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CO20_1 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 37.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11 or SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 37;

(c) fragments of the amino acid sequence of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 37.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 176 to nucleotide 520;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 317 to nucleotide 520;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 118 to nucleotide 413;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO223_3 deposited under accession number ATCC 98291;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO223_3 deposited under accession number ATCC 98291;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:14 from nucleotide 176 to nucleotide 520; the nucleotide sequence of SEQ ID NO:14 from nucleotide 317 to nucleotide 520; the nucleotide sequence of SEQ ID NO:14 from nucleotide 118 to nucleotide 413; the nucleotide sequence of the full-length protein coding sequence of clone CO223_3 deposited under accession number ATCC 98291; or the nucleotide sequence of the mature protein coding sequence of clone CO223_3 deposited under accession number ATCC 98291. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 80.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:14.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 80;

(c) fragments of the amino acid sequence of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:15 or the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 80.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 303 to nucleotide 542;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 1 to nucleotide 435;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO310_2 deposited under accession number ATCC 98279;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO310_2 deposited under accession number ATCC 98279;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:16 from nucleotide 303 to nucleotide 542; the nucleotide sequence of SEQ ID NO:16 from nucleotide 1 to nucleotide 435; the nucleotide sequence of the full-length protein coding sequence of clone CO310_2 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CO310_2 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 44.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:16.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:17;

(b) the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 44;

(c) fragments of the amino acid sequence of SEQ ID NO:17; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:17 or the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 44.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 40 to nucleotide 455;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 85 to nucleotide 455;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 265 to nucleotide 515;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CP258_3 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CP258_3 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:18 from nucleotide 40 to nucleotide 455; the nucleotide sequence of SEQ ID NO:18 from nucleotide 85 to nucleotide 455; the nucleotide sequence of SEQ ID NO:18 from nucleotide 265 to nucleotide 515; the nucleotide sequence of the full-length protein coding sequence of clone CP258_3 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CP258_3 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19 from amino acid 64 to amino acid 138.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:18.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 64 to amino acid 138;

(c) fragments of the amino acid sequence of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:19 or the amino acid sequence of SEQ ID NO:19 from amino acid 64 to amino acid 138.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 105 to nucleotide 1007;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 801 to nucleotide 1007;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 352;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CW1155_3 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CW1155_3 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:20 from nucleotide 105 to nucleotide 1007; the nucleotide sequence of SEQ ID NO:20 from nucleotide 801 to nucleotide 1007; the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 352; the nucleotide sequence of the full-length protein coding

sequence of clone CW1155_3 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CW1155_3 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 83.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:20.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 83;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:21 or the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 83.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 11 to nucleotide 1699;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 1682 to nucleotide 1699;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 737 to nucleotide 1134;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CZ247_2 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CZ247_2 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:22 from nucleotide 11 to nucleotide 1699; the nucleotide sequence of SEQ ID NO:22 from nucleotide 1682 to nucleotide 1699; the nucleotide sequence of SEQ ID NO:22 from nucleotide 737 to nucleotide 1134; the nucleotide sequence of the full-length protein coding sequence of clone CZ247_2 deposited under accession number ATCC 98279; or the nucleotide sequence of the mature protein coding sequence of clone CZ247_2 deposited under accession number ATCC 98279. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23 from amino acid 298 to amino acid 374.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:22.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 298 to amino acid 374;

(c) fragments of the amino acid sequence of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:23 or the amino acid sequence of SEQ ID NO:23 from amino acid 298 to amino acid 374.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "CB107_1"

A polynucleotide of the present invention has been identified as clone "CB107_1". CB107_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CB107_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CB107_1 protein").

The nucleotide sequence of the 5' portion of CB107_1 as presently determined is reported in SEQ ID NO:1. An additional internal nucleotide sequence from CB107_1 as presently determined is reported in SEQ ID NO:2. What applicants believe is the proper

reading frame and the predicted amino acid sequence encoded by such internal sequence is reported in SEQ ID NO:3. Additional nucleotide sequence from the 3' portion of CB107_1, including the polyA tail, is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CB107_1 should be approximately 3300 bp.

The nucleotide sequence disclosed herein for CB107_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CB107_1 demonstrated at least some similarity with sequences identified as AA121485 (zn80a02.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone 564458 3'), AA428192 (zw51b08.s1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone 773559 3'), D83018 (Human mRNA for nel-related protein 2, complete cds), F10919 (H. sapiens partial cDNA sequence; clone c-3lg01), H15375 (ym28d09.r1 Homo sapiens cDNA clone 49527 5' similar to SP A54105 A54105 FIBRILLIN-2 PRECURSOR), U48245 (Rattus norvegicus protein kinase C-binding protein Nel mRNA, complete cds), U59230 (Mus musculus mel (MEL91) mRNA, complete cds), and W28387 (46c5 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA). The predicted amino acid sequence disclosed herein for CB107_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CB107_1 protein demonstrated at least some similarity to sequences identified as D83018 (nel-related protein 2 [Homo sapiens]), R05222 (Antigen GX5401FL encoded by Eimeria tenella genomic DNA), R79964 (Connective tissue growth factor), U48245 (RNU48245_1 protein kinase C-binding protein Nel [Rattus norvegicus]), and U59230 (mel [Mus musculus]). Based upon sequence similarity, CB107_1 proteins and each similar protein or peptide may share at least some activity.

Clone "CG300_3"

A polynucleotide of the present invention has been identified as clone "CG300_3". CG300_3 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CG300_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CG300_3 protein").

The nucleotide sequence of CG300_3 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CG300_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6. Amino acids 30 to 42 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 43, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CG300_3 should be approximately 1800 bp.

The nucleotide sequence disclosed herein for CG300_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CG300_3 demonstrated at least some similarity with sequences identified as N40185 (yy44d08.s1 Homo sapiens cDNA clone 276399 3') and W01791 (za72d06.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 298091 5'). Based upon sequence similarity, CG300_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts four potential transmembrane domains within the CG300_3 protein sequence, centered around amino acids 34, 98, 151, and 179 of SEQ ID NO:6, respectively.

Clone "CJ145_1"

A polynucleotide of the present invention has been identified as clone "CJ145_1". CJ145_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CJ145_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CJ145_1 protein").

The nucleotide sequence of CJ145_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CJ145_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 6 to 18 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 19, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CJ145_1 should be approximately 3600 bp.

The nucleotide sequence disclosed herein for CJ145_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CJ145_1 demonstrated at least some similarity with sequences identified as R43655 (yc86b04.s1 Homo sapiens cDNA clone 22829 3'), R50995 (yg63f06.s1 Homo sapiens cDNA clone 37377 3' similar to contains MER22 repetitive element), and W92748 (zd92h03.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 356981 3'). Based upon sequence similarity, CJ145_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of CJ145_1 indicates that it may contain a CA simple repeat element.

Clone "CJ160_11"

A polynucleotide of the present invention has been identified as clone "CJ160_11". CJ160_11 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CJ160_11 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CJ160_11 protein").

The nucleotide sequence of CJ160_11 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CJ160_11 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CJ160_11 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for CJ160_11 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CJ160_11 demonstrated at least some similarity with sequences identified as AA024511 (ze76e04.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 364926 3') and AC000074 (00884; HTGS phase 3, complete sequence). Based upon sequence similarity, CJ160_11 proteins and each similar protein or peptide may share at least some activity.

Clone "CO20_1"

A polynucleotide of the present invention has been identified as clone "CO20_1". CO20_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO20_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO20_1 protein").

The nucleotide sequence of the 5' portion of CO20_1 as presently determined is reported in SEQ ID NO:11. What applicants presently believe is the proper reading frame for the coding region is indicated in SEQ ID NO:12. The predicted amino acid sequence of the CO20_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain. Additional nucleotide sequence from the 3' portion of CO20_1, including the polyA tail, is reported in SEQ ID NO:13.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO20_1 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for CO20_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO20_1 demonstrated at least some similarity with sequences identified as AA045770 (zl68b10.s1 Stratagene colon (#937204) Homo sapiens cDNA clone 509755 3' similar to SW:R13A_HUMAN P40429 60S RIBOSOMAL PROTEIN L13A), AA070899 (zm66c01.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone 530592 3' similar to contains Alu repetitive element), AA325205 (EST28155 Cerebellum II Homo sapiens cDNA 5' end), N22253 (yw36a08.s1 Homo sapiens cDNA clone 254294 3' similar to SP S29539 S29539 BASIC PROTEIN, 23K), R01933 (ye85g07.s1 Homo sapiens cDNA clone 124572 3' similar to SP:S29539 S29539 BASIC PROTEIN, 23K), R12008 (yf51f04.r1 Homo sapiens cDNA clone 25456 5'), R39848 (yf51f04.s1 Homo sapiens cDNA clone 25456 3' similar to contains Alu repetitive element;contains PTR5 repetitive element), R56565 (yg91c12.r1 Homo sapiens cDNA clone 40891 5'), T19487 (Human gene signature HUMGS00543), T30988 (EST25695 Homo sapiens cDNA 5' end similar to None), U37026 (Rattus norvegicus brain sodium channel beta 2 subunit (SCNB2) mRNA, complete cds), and X56932 (H.sapiens mRNA for 23 kD highly basic protein). The predicted amino acid

sequence disclosed herein for CO20_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CO20_1 protein demonstrated at least some similarity to sequences identified as U37026 (sodium channel beta 2 subunit [Rattus norvegicus]), U58658 (unknown [Homo sapiens]), and X56932 (23 kD highly basic protein [Homo sapiens]). The sodium channel beta 2 subunit is a glycoprotein with an extracellular domain containing an immunoglobulin-like fold with similarity to the neural cell adhesion molecule contactin. Based upon sequence similarity, CO20_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of CO20_1 indicates that it may contain an Alu repetitive element.

Clone "CO223_3"

A polynucleotide of the present invention has been identified as clone "CO223_3". CO223_3 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO223_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO223_3 protein").

The nucleotide sequence of CO223_3 as presently determined is reported in SEQ ID NO:14. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CO223_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:15. Amino acids 35 to 47 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 48, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO223_3 should be approximately 700 bp.

The nucleotide sequence disclosed herein for CO223_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CO223_3 demonstrated at least some similarity with sequences identified as AA004498 (zh87b06.r1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 428243 5' similar to gb M62505 C5A ANAPHYLATOXIN CHEMOTACTIC RECEPTOR (HUMAN);contains L1.t1 L1 repetitive element) and U47924 (Human chromosome 12p13 gene cluster, surface antigen CD4 (CD4), A, B, G-protein beta-3

subunit (GNB3), isopeptidase T (ISOT) and triosephosphate isomerase (TPI) genes, complete cds). Based upon sequence similarity, CO223_3 proteins and each similar protein or peptide may share at least some activity.

The 3' end of the CO223_3 polynucleotide sequence contains a 54-bp sequence that is repeated three times in the clone; these repeats begin at positions 314, 368, and 422 of SEQ ID NO:14 and encode amino acids 47 to 64, 65 to 82, and 83 to 99 of SEQ ID NO:15, respectively.

Clone "CO310_2"

A polynucleotide of the present invention has been identified as clone "CO310_2". CO310_2 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CO310_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CO310_2 protein").

The nucleotide sequence of CO310_2 as presently determined is reported in SEQ ID NO:16. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CO310_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:17.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CO310_2 should be approximately 1400 bp.

The nucleotide sequence disclosed herein for CO310_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. No hits were found in the database. The nucleotide sequence of CO310_2 indicates that it may contain an L1 repetitive element.

Clone "CP258_3"

A polynucleotide of the present invention has been identified as clone "CP258_3". CP258_3 was isolated from a human adult salivary gland cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CP258_3 is a full-length clone,

including the entire coding sequence of a secreted protein (also referred to herein as "CP258_3 protein").

The nucleotide sequence of CP258_3 as presently determined is reported in SEQ ID NO:18. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CP258_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:19. Amino acids 3 to 15 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 16, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CP258_3 should be approximately 560 bp.

The nucleotide sequence disclosed herein for CP258_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. No hits were found in the database.

Clone "CW1155_3"

A polynucleotide of the present invention has been identified as clone "CW1155_3". CW1155_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CW1155_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CW1155_3 protein").

The nucleotide sequence of CW1155_3 as presently determined is reported in SEQ ID NO:20. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CW1155_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:21. Amino acids 220 to 232 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 233, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CW1155_3 should be approximately 1170 bp.

The nucleotide sequence disclosed herein for CW1155_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CW1155_3 demonstrated at least some similarity with sequences identified as AA169043 (ms36h08.r1 Stratagene mouse heart (#937316) Mus musculus

cDNA clone 613695 5'), D86145 (Rat mRNA), and H29261 (ym32b03.s1 Homo sapiens cDNA clone 49733 3'). Based upon sequence similarity, CW1155_3 proteins and each similar protein or peptide may share at least some activity.

Clone "CZ247_2"

A polynucleotide of the present invention has been identified as clone "CZ247_2". CZ247_2 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. CZ247_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "CZ247_2 protein").

The nucleotide sequence of CZ247_2 as presently determined is reported in SEQ ID NO:22. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the CZ247_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:23. Amino acids 545 to 557 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 558, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone CZ247_2 should be approximately 2300 bp.

The nucleotide sequence disclosed herein for CZ247_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. CZ247_2 demonstrated at least some similarity with sequences identified as T09256 (Human ara Kb beta-galactosidase fusion protein coding sequence), W27222 (26h9 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA), and W72736 (zd71e02.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 346106 3'). The predicted amino acid sequence disclosed herein for CZ247_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted CZ247_2 protein demonstrated at least some similarity to sequences identified as R88069 (Human ara Kb beta-galactosidase fusion protein). Based upon sequence similarity, CZ247_2 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones CB107_1, CG300_3, CJ145_1, CJ160_11, CO20_1, CO223_1, CO310_2, CP258_3, CW1155_3 and CZ247_2 were deposited on December 17, 1996 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98279, from which each clone comprising a particular polynucleotide is obtainable. Clone CO223_3 was deposited on January 9, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98291. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* **19**: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* **9**: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
CB107_1	SEQ ID NO:24
CG300_3	SEQ ID NO:25
CJ145_1	SEQ ID NO:26
CJ160_11	SEQ ID NO:27
CO20_1	SEQ ID NO:28
CO223_3	SEQ ID NO:29
CO310_2	SEQ ID NO:30
CP258_3	SEQ ID NO:31
CW1155_3	SEQ ID NO:32
CZ247_2	SEQ ID NO:33

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in

fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the

polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or

polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides .

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which

the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays

for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein

and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines

or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium

Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where

abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect

the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 33

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATCGCACCT GTCCAACTTG CAATGACTTC CATGGACTTG TGCAGAAAAT CATGGAGCTA 60

CAGGATATTT TAGCCAAAAC ATCAGCCAAG CTGTCTCGAG CTGAACAGCG AATGAATAGA 120

TTGGATCAGT GCTATTGTGA AAGGACTTGC ACCATGAAGG GAACCACCTA CCGAGAATTT 180

GAGTCCTGGA TAGACGGCTG TAAGAACTGC ACATGCCTGA ATGGAACCAT CCAGTGTGAA 240

ACTCTAATCT GCCCAAATCC TGACTGCCCA CTTAAGTCCG CTCTTGCGTA TGTGGATGGC 300

AAATGCTGTA AGGAATGCAA ATCGATATTC CAATTTCAAG GACGAACCTA CTTTGAAGGA 360

GAAAGAAATA CA 372

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 761 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTCTGGAGTA TGTGTTCTCT ATGAGTGCAA GGACCAGACC ATGAAATTTG TTGAGAGTTC 60

AGGCTGTCCA GCTTTGGATT GTCCAGAGTC TCATCAGATA ACCTTGTNTC ACAGCTGTTG 120

CAAAGTTTGT AAAGGTTATG ATTTTTGTTT TGAAAGGCAT AACTGCATGG AGAATTCCAT 180

CTGCAGAAAT NTGAATGACA GGGCTGTTTG TAGCTGTCGA GATGGTTTTA GGGTTTTTCG 240

AGAGGATAAT GCCTACTGTG AAGACATNGA TGAGTGTGCT GAAGGGCGCC ATTACTGTNG 300

TGAAAATACA ATGTGTGTCA ACACCCCGGG TTCTTTTATG TGCATCTGCA AAACTGGATA 360

CATCAGAATT GATGATTATT CATGTACAGA ACATGATGAG TGTATCACAA ATCAGCACAG 420

CTGTGATGAA AATGCTTTAT GCTTCAACAC TGTTGGAGGA CACAACTGTG TTTGCAAGCC 480

GGGCTATACA GGGAATGGAA CGACATGCAA AGCATTTTGC AAAGATGGCT GTAGGAATGG 540

AGGAGCCTGT ATTGCCGCTA ATGTGTGTGC CTGCCCACAA GGCTTCACTG GACCCAGCTG 600

TGAAACGGAC ATTGATGAAT GCTCTGATGG TTTTGTTCAA TGTGACAGTC GTGCTAATTG 660

CATTAACCTG CCTGGATGGT ACCACTGTGA GTGCAGAGAT GGCTACCATG ACAATGGGAT 720

GTTTTCACCA AGTGGAGAAT CGTGTGAAGA TATTGATGAG T 761

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Lys Phe Val Glu Ser Ser Gly Cys Pro Ala Leu Asp Cys Pro Glu
1 5 10 15

Ser His Gln Ile Thr Leu Xaa His Ser Cys Cys Lys Val Cys Lys Gly
20 25 30

Tyr Asp Phe Cys Phe Glu Arg His Asn Cys Met Glu Asn Ser Ile Cys
35 40 45

Arg Asn Xaa Asn Asp Arg Ala Val Cys Ser Cys Arg Asp Gly Phe Arg
50 55 60

Val Phe Arg Glu Asp Asn Ala Tyr Cys Glu Asp Xaa Asp Glu Cys Ala
65 70 75 80

Glu Gly Arg His Tyr Cys Xaa Glu Asn Thr Met Cys Val Asn Thr Pro
85 90 95

Gly Ser Phe Met Cys Ile Cys Lys Thr Gly Tyr Ile Arg Ile Asp Asp
100 105 110

Tyr Ser Cys Thr Glu His Asp Glu Cys Ile Thr Asn Gln His Ser Cys
115 120 125

Asp Glu Asn Ala Leu Cys Phe Asn Thr Val Gly Gly His Asn Cys Val
130 135 140

Cys Lys Pro Gly Tyr Thr Gly Asn Gly Thr Thr Cys Lys Ala Phe Cys
145 150 155 160

Lys Asp Gly Cys Arg Asn Gly Gly Ala Cys Ile Ala Ala Asn Val Cys
165 170 175

Ala Cys Pro Gln Gly Phe Thr Gly Pro Ser Cys Glu Thr Asp Ile Asp

180 185 190

Glu Cys Ser Asp Gly Phe Val Gln Cys Asp Ser Arg Ala Asn Cys Ile
195 200 205

Asn Leu Pro Gly Trp Tyr His Cys Glu Cys Arg Asp Gly Tyr His Asp
210 215 220

Asn Gly Met Phe Ser Pro Ser Gly Glu Ser Cys Glu Asp Ile Asp Glu
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 342 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCAGAAAATT TTCCTCTAGA TCAGAATCTT CAAGAATCAG TTAGGTTCCT CACTGCAAGA 60

AATAAAATGT CAGGCAGTGA ATGAATTATA TTTTAAGAAG TAAAGCAAAG AAGCTATAAC 120

ATGTTATGTA CAGTACACTC TGAAAAGAAA TCTGAAACAA GTTATTGTAA TGATAAAAAT 180

AATGCACAGG CATGGTTACT TAATATTTTC TAACAGGAAA AGTCATCCCT ATTTCCTTGT 240

TTTACTGCAC TTAATATTAT TTGGTTGAAT TTGTTCAGTA TAAGTTCGTT CCTTGTGCAA 300

AATTAAATAA ATATTTTTCT TACCTTAAAA AAAAAAAAAA AA 342

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1445 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GTGCGCATGG GGACGCTATA GCAATTCGTT TGCTGTCCTT CCTCTCCTTC GAAGATGACA 60

AGGCCTACCA TCGTTTCTTC CTGCCTTTGG GCCGTCAGGC AGTTGGTTGG GACCCGCTCC 120

AACCCTCGGT TCTTCCTGCA ATACAGTGGA TACAATTTGT CATGGCTACT CTGAGATAAG 180

ACCACTTTTT TATCTGAGCT TCTGTGACCT GCTCCTGGGA CTTTGCTGGC TCACGGAGAC 240

ACTTCTCTAT GGAGCTTCAG TAGCAAATAA GGACATCATC TGCTATAACC TACAAGCAGT 300

TGGACAGATA TTCTACATTT CCTCATTTCT CTACACCGTC AATTACATCT GGTATTTGTA 360

CACAGAGCTG AGGATGAAAC ACACCCAAAG TGGACAGAGC ACATCTCCAC TGGTGATAGA 420

TTATACTTGT CGATTTTGTC AAATGGCCTT TGTTTTCTCA AGCCTGATAC CTCTGCTATT 480

GATGACACCT GTATTCTGTC TGGGAAATAC TAGTGAATGT TTCCAAAACT TCAGTCAGAG 540

CCACAATTGT ATCTTGATGC ACTCACCACC ATCAGCCATG GCTGAACTTC CACCTTCTGC 600

CAACACATCT GTCTGTAGCA CACTTTATTT TTATGGTATC GCCATTTTCC TGGGCAGCTT 660

TGTACTCAGC CTCCTTACCA TTATGGTCTT ACTTATCCGA GCCCAGACAT TGTATAAGAA 720

GTTTGTGAAG TCAACTGGCT TTCTGGGGAG TGAACAGTGG GCAGTGATTC ACATTGTGGA 780

CCAACGGGTG CGCTTCTACC CAGTGGCCTT CTTTTGCTGC TGGGGCCCAG CTGTCATTCT 840

AATGATCATA AAGCTGACTA AGCCACAGGA CACCAAGCTT CACATGGCCC TTTATGTTCT 900

CCAGGCTCTA ACGGCAACAT CTCAGGGTCT ACTCAACTGT GGAGTATATG GCTGGACGCA 960

GCACAAATTC CACCAACTAA AGCAGGAGGC TCGGCGTGAT GCAGATACCC AGACACCATT 1020

ATTATGCTCA CAGAAGAGAT TCTATAGCAG GGGCTTAAAT TCACTGGAAT CCACCCTGAC 1080

TTTTCCTGCC AGTACTTCTA CCATTTTTTG AAACTACAAT ACTGGAACAT CCAGGAACTG 1140

GAGTTATTCT ACGCTAATGG ATTGGAAAGA ATGTTGGGAA AGGACATCTT AAATCTTTTC 1200

TAACTATGCC CTAAACTGCA GAACTCAAAG GAAATATAGT GCCATTGTTA GTAGTCATTC 1260

TAGATGAATT GGGAGTATCT CTCCAGTTAT TCCCAGATTC ACTAGTGATC CTTAAAGTCT 1320

CTATTCAGGG AGAGGAAGAC ACTTTCCATC TCAGAGATAG ACTCGTGTTA CCTTGATGGA 1380

TATTGGATTT GTCTAAGTCT CTTCTAGAAA AAATAAATTC TAGATTATTA AAAAAAAAAA 1440

AAAAA 1445

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys His Thr Gln Ser Gly Gln Ser Thr Ser Pro Leu Val Ile Asp
1 5 10 15

Tyr Thr Cys Arg Phe Cys Gln Met Ala Phe Val Phe Ser Ser Leu Ile
20 25 30

Pro Leu Leu Leu Met Thr Pro Val Phe Cys Leu Gly Asn Thr Ser Glu
35 40 45

Cys Phe Gln Asn Phe Ser Gln Ser His Asn Cys Ile Leu Met His Ser
50 55 60

Pro Pro Ser Ala Met Ala Glu Leu Pro Pro Ser Ala Asn Thr Ser Val
65 70 75 80

Cys Ser Thr Leu Tyr Phe Tyr Gly Ile Ala Ile Phe Leu Gly Ser Phe
85 90 95

Val Leu Ser Leu Leu Thr Ile Met Val Leu Leu Ile Arg Ala Gln Thr
100 105 110

Leu Tyr Lys Lys Phe Val Lys Ser Thr Gly Phe Leu Gly Ser Glu Gln
115 120 125

Trp Ala Val Ile His Ile Val Asp Gln Arg Val Arg Phe Tyr Pro Val
130 135 140

Ala Phe Phe Cys Cys Trp Gly Pro Ala Val Ile Leu Met Ile Ile Lys
145 150 155 160

Leu Thr Lys Pro Gln Asp Thr Lys Leu His Met Ala Leu Tyr Val Leu
165 170 175

Gln Ala Leu Thr Ala Thr Ser Gln Gly Leu Leu Asn Cys Gly Val Tyr
180 185 190

Gly Trp Thr Gln His Lys Phe His Gln Leu Lys Gln Glu Ala Arg Arg
195 200 205

Asp Ala Asp Thr Gln Thr Pro Leu Leu Cys Ser Gln Lys Arg Phe Tyr
210 215 220

Ser Arg Gly Leu Asn Ser Leu Glu Ser Thr Leu Thr Phe Pro Ala Ser
225 230 235 240

Thr Ser Thr Ile Phe
245

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3550 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CCCCTCGATA ATGGATTACT AAATGGGATA CACGCTGTAC CAGTTCGCTC CGAGCCCCGG 60

CCGCCTGTCC GTCGATGCAC CGAAAAGGGT GAAGTAGAGA AATAAAGTCT CCCCGCTGAA 120

CTACTATGAG GTCAGAAGCC TTGCTCCTAT ATTTCACACT GCTACACTTT GCTGGGGCTG 180

GTTTCCCAGA AGATTCTGAG CCAATCAGTA TTTCGCATGG CAACTATACA AAACAGTATC 240

CGGTGTTTGT GGGCCACAAG CCAGGACGGA ACACCACACA GAGGCACAGG CTGGACATCC 300

AGATGATTAT GATCATGAAC GGAACCCTCT ACATTGCTGC TAGGGACCAT ATTTATACTG 360

TTGATATAGA CACATCACAC ACSGAAGAAA TTTATTGTAG CAAAAAACTG ACATGGAAAT 420

CTAGACAGGC CGATGTAGAC ACATGCAGAA TGAAGGGAAA ACATAAGGAT GAGTGCCACA 480

ACTTTATTAA AGTTCTTCTA AAGAAAAACG ATGATGCATT GTTTGTCTGT GGAACTAATG 540

CCTTCAACCC TTCCTGCAGA AACTATAAGA TGGATACATT GGAACCATTC GGGGATGAAT 600

TCAGCGGAAT GGCCAGATGC CCATATGATG CCAAACATGC CAACGTTGCA CTGTTTGCAG 660

ATGGAAAACT ATACTCAGCC ACAGTGACTG ACTTCCTTGC CATTGACGCA GTCATTTACC 720

GGAGTCTTGG AGAAAGCCCT ACCCTGCGGA CCGTCAAGCA CGATTCAAAA TGGTTGAAAG 780

AACCATACTT TGTTCAAGCC GTGGATTACG GAGATTATAT CTACTTCTTC TTCAGGGAAA 840

TAGCAGTGGA GTATAACACC ATGGGAAAGG TAGTTTTCCC AAGAGTGGCT CAGGTTTGTA 900

AGAATGATAT GGGAGGATCT CAAAGAGTCC TGGAGAAACA GTGGACGTCG TTCCTGAAGG 960

CGCGCTTGAA CTGCTCAGTT CCTGGAGACT CTCATTTTTA TTTCAACATT CTCCAGGCAG 1020

TTACAGATGT GATTCGTATC AACGGGCGTG ATGTTGTCCT GGCAACGTTT TCTACACCTT 1080

ATAACAGCAT CCCTGGGTCT GCAGTCTGTG CCTATGACAT GCTTGACATT GCCAGTGTTT 1140

TTACTGGGAG ATTCAAGGAA CAGAAGTCTC CTGATTCCAC CTGGACACCA GTTCCTGATG 1200

AACGAGTTCC TAAGCCCAGG CCAGGTTGCT GTGCTGGCTC ATCCTCCTTA GAAAGATATG 1260

CAACCTCCAA TGAGTTCCCT GATGATACCC TGAACTTCAT CAAGACGCAC CCGCTCATGG 1320

ATGAGGCAGT GCCCTCCATC TTCAACAGGC CATGGTTCCT GAGAACAATG GTCAGATACC 1380

GCCTTACCAA AATTGCAGTG GACACAGCTG CTGGGCCATA TCAGAATCAC ACTGTGGTTT 1440

TTCTGGGATC AGAGAAGGGA ATCATCTTGA AGTTTTTGGC CAGAATAGGA AATAGTGGTT 1500

TTCTAAATGA CAGCCTTTTC CTGGAGGAGA TGAGTGTTTA CAACTCTGAA AAATGCAGCT 1560

ATGATGGAGT CGAAGACAAA AGGATCATGG GCATGCAGCT GGACAGAGCA AGCAGCTCTC 1620

TGTATGTTGC GTTCTCTACC TGTGTGATAA AGGTTCCCCT TGGCCGGTGT GAACGACATG 1680

GGAAGTGTAA AAAAACCTGT ATTGCCTCCA GAGACCCATA TTGTGGATGG ATAAAGGAAG 1740

GTGGTGCCTG CAGCCATTTA TCACCCAACA GCAGACTGAC TTTTGAGCAG GACATAGAGC 1800

GTGGCAATAC AGATGGTCTG GGGGACTGTC ACAATTCCTT TGTGGCACTG AATGGAGTGA 1860

TTCGGGAAAG TTACCTCAAA GGCCACGACC AGCTGGTTCC CGTCACCCTC TTGGCCATTG 1920

CAGTCATCCT GGCTTTCGTC ATGGGGGCCG TCTTCTCGGG CATCACCGTC TACTGCGTCT 1980

GTGATCATCG GCGCAAAGAC GTGGCTGTGG TGCAGCGCAA GGAGAAGGAG CTCACCCACT 2040

CGCGCCGGGG CTCCATGAGC AGCGTCACCA AGCTCAGCGG CCTCTTTGGG GACACTCAAT 2100

CCAAAGACCC AAAGCCGGAG GCCATCCTCA CGCCACTCAT GCACAACGGC AAGCTCGCCA 2160

CTCCCGGCAA CACGGCCAAG ATGCTCATTA AAGCAGACCA GCACCACCTG GACCTGACGG 2220

CCCTCCCCAC CCCAGAGTCA ACCCCAACGC TGCAGCAGAA GCGGAAGCCC AGCCGCGGCA 2280

GCCGCGAGTG GGAGAGGAAC CAGAACCTCA TCAATGCCTG CACAAAGGAC ATGCCCCCCA 2340

TGGGCTCCCC TGTGATTCCC ACGGACCTGC CCCTGCGGGC CTCCCCCAGC CACATCCCCA 2400

GCGTGGTGGT CCTGCCCATC ACGCAGCAGG GCTACCAGCA TGAGTACGTG GACCAGCCCA 2460

AAATGAGCGA GGTGGCCCAG ATGGCGCTGG AGGACCAGGC CGCCACACTG GAGTATAAGA 2520

CCATCAAGGA ACATTTCAGC AGCAAGAGTC CCAACCATGG GGTGAACCTT GTGGAGAACC 2580

TGGACAGCCT GCCCCCCAAA GTTCCACAGC GGGAGGCCTC CCTGGGTCCC CCGGGAGCCT 2640

CCCTGTTTCA GACCGGTTTA AGCAAGCGGC TGGAAATGCA CCACTCCTTT TCCTACGGGG 2700

TTGACTATAA GAGGAGCTAC CCCACGAACT CGCTCACGAG AAGCCACCAG GCCACCACTC 2760

TCAAAAGAAA CAACACTAAC TCCTCCAATT CCTCTCACCT CTCCAGAAAC CAGAGCTTTG 2820

GCAGGGGAGA CAACCCGCCG CCCGCCCCGC AGAGGGTGGA CTCCATCCAG GTGCACAGCT 2880

CCCAGCCATC TGGCCAGGCC GTGACTGTCT CGAGGCAGCC CAGCCTCAAC GCCTACAACT 2940

CACTGACAAG GTCGGGGCTG AAGCGTACGC CCTCGCTAAA GCCGGACGTA CCCCCCAAAC 3000

CATCCTTTGC TCCCCTTTCC ACATCCATGA AGCCCAATGA TGCGTGTACA TAATCCCAGG 3060

GGGAGGGGGT CAGGTGTCGA ACCAGCAGGC AAGGCGAGGT GCCCGCTCAG CTCAGCAAGG 3120

TTCTCAACTG CCTCGAGTAC CCACCAGACC AAGAAGGCCT GCGGCAGAGC CGAGGACGCT 3180

GGGTCCTCCT CTCTGGGACA CAGGGGTACT CACGAAAACT GGGCCGCGTG GTTTGGTGAA 3240

GGTTTGCAAC GGCGGGGACT CACCTTCATT CTCTTCCTTC ACTTTCCCCC ACACCCTACA 3300

ACAGGTCGGA CCCACAAAAG ACTTCAGTTA TCATCACAAA CATGAGCCAA AAGCACATAC 3360

ATACCCCATC CCCCACCCCC ACACACACAC ACACATGCAC ACAACACATA CACACACACG 3420

CACAGAGGTG AACAGAAACT GAAACATTTT GTCCACAACT TCACGGGACG TGGCCAGACT 3480

GGGTTTGCGT TCCAACCTGC AAAACACAAA TACATTTTTT AAAATCAAGA AAATTTAAAA 3540

AAAAAAAAAA 3550

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 975 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Arg Ser Glu Ala Leu Leu Leu Tyr Phe Thr Leu Leu His Phe Ala
1 5 10 15

Gly Ala Gly Phe Pro Glu Asp Ser Glu Pro Ile Ser Ile Ser His Gly
20 25 30

Asn Tyr Thr Lys Gln Tyr Pro Val Phe Val Gly His Lys Pro Gly Arg
35 40 45

Asn Thr Thr Gln Arg His Arg Leu Asp Ile Gln Met Ile Met Ile Met
50 55 60

Asn Gly Thr Leu Tyr Ile Ala Ala Arg Asp His Ile Tyr Thr Val Asp
65 70 75 80

Ile Asp Thr Ser His Thr Glu Glu Ile Tyr Cys Ser Lys Lys Leu Thr

85 90 95

Trp Lys Ser Arg Gln Ala Asp Val Asp Thr Cys Arg Met Lys Gly Lys
100 105 110

His Lys Asp Glu Cys His Asn Phe Ile Lys Val Leu Leu Lys Lys Asn
115 120 125

Asp Asp Ala Leu Phe Val Cys Gly Thr Asn Ala Phe Asn Pro Ser Cys
130 135 140

Arg Asn Tyr Lys Met Asp Thr Leu Glu Pro Phe Gly Asp Glu Phe Ser
145 150 155 160

Gly Met Ala Arg Cys Pro Tyr Asp Ala Lys His Ala Asn Val Ala Leu
165 170 175

Phe Ala Asp Gly Lys Leu Tyr Ser Ala Thr Val Thr Asp Phe Leu Ala
180 185 190

Ile Asp Ala Val Ile Tyr Arg Ser Leu Gly Glu Ser Pro Thr Leu Arg
195 200 205

Thr Val Lys His Asp Ser Lys Trp Leu Lys Glu Pro Tyr Phe Val Gln
210 215 220

Ala Val Asp Tyr Gly Asp Tyr Ile Tyr Phe Phe Phe Arg Glu Ile Ala
225 230 235 240

Val Glu Tyr Asn Thr Met Gly Lys Val Val Phe Pro Arg Val Ala Gln
245 250 255

Val Cys Lys Asn Asp Met Gly Gly Ser Gln Arg Val Leu Glu Lys Gln
260 265 270

Trp Thr Ser Phe Leu Lys Ala Arg Leu Asn Cys Ser Val Pro Gly Asp
275 280 285

Ser His Phe Tyr Phe Asn Ile Leu Gln Ala Val Thr Asp Val Ile Arg
290 295 300

Ile Asn Gly Arg Asp Val Val Leu Ala Thr Phe Ser Thr Pro Tyr Asn
305 310 315 320

Ser Ile Pro Gly Ser Ala Val Cys Ala Tyr Asp Met Leu Asp Ile Ala
325 330 335

Ser Val Phe Thr Gly Arg Phe Lys Glu Gln Lys Ser Pro Asp Ser Thr
340 345 350

Trp Thr Pro Val Pro Asp Glu Arg Val Pro Lys Pro Arg Pro Gly Cys
355 360 365

Cys Ala Gly Ser Ser Ser Leu Glu Arg Tyr Ala Thr Ser Asn Glu Phe
370 375 380

Pro Asp Asp Thr Leu Asn Phe Ile Lys Thr His Pro Leu Met Asp Glu
385 390 395 400

Ala Val Pro Ser Ile Phe Asn Arg Pro Trp Phe Leu Arg Thr Met Val
405 410 415

Arg Tyr Arg Leu Thr Lys Ile Ala Val Asp Thr Ala Ala Gly Pro Tyr
420 425 430

Gln Asn His Thr Val Val Phe Leu Gly Ser Glu Lys Gly Ile Ile Leu
435 440 445

Lys Phe Leu Ala Arg Ile Gly Asn Ser Gly Phe Leu Asn Asp Ser Leu
450 455 460

Phe Leu Glu Glu Met Ser Val Tyr Asn Ser Glu Lys Cys Ser Tyr Asp
465 470 475 480

Gly Val Glu Asp Lys Arg Ile Met Gly Met Gln Leu Asp Arg Ala Ser
485 490 495

Ser Ser Leu Tyr Val Ala Phe Ser Thr Cys Val Ile Lys Val Pro Leu
500 505 510

Gly Arg Cys Glu Arg His Gly Lys Cys Lys Lys Thr Cys Ile Ala Ser
515 520 525

Arg Asp Pro Tyr Cys Gly Trp Ile Lys Glu Gly Gly Ala Cys Ser His
530 535 540

Leu Ser Pro Asn Ser Arg Leu Thr Phe Glu Gln Asp Ile Glu Arg Gly
545 550 555 560

Asn Thr Asp Gly Leu Gly Asp Cys His Asn Ser Phe Val Ala Leu Asn
565 570 575

Gly Val Ile Arg Glu Ser Tyr Leu Lys Gly His Asp Gln Leu Val Pro
580 585 590

Val Thr Leu Leu Ala Ile Ala Val Ile Leu Ala Phe Val Met Gly Ala
595 600 605

Val Phe Ser Gly Ile Thr Val Tyr Cys Val Cys Asp His Arg Arg Lys
610 615 620

Asp Val Ala Val Val Gln Arg Lys Glu Lys Glu Leu Thr His Ser Arg
625 630 635 640

Arg Gly Ser Met Ser Ser Val Thr Lys Leu Ser Gly Leu Phe Gly Asp
645 650 655

Thr Gln Ser Lys Asp Pro Lys Pro Glu Ala Ile Leu Thr Pro Leu Met
660 665 670

His Asn Gly Lys Leu Ala Thr Pro Gly Asn Thr Ala Lys Met Leu Ile

675 680 685

Lys Ala Asp Gln His His Leu Asp Leu Thr Ala Leu Pro Thr Pro Glu
690 695 700

Ser Thr Pro Thr Leu Gln Gln Lys Arg Lys Pro Ser Arg Gly Ser Arg
705 710 715 720

Glu Trp Glu Arg Asn Gln Asn Leu Ile Asn Ala Cys Thr Lys Asp Met
725 730 735

Pro Pro Met Gly Ser Pro Val Ile Pro Thr Asp Leu Pro Leu Arg Ala
740 745 750

Ser Pro Ser His Ile Pro Ser Val Val Val Leu Pro Ile Thr Gln Gln
755 760 765

Gly Tyr Gln His Glu Tyr Val Asp Gln Pro Lys Met Ser Glu Val Ala
770 775 780

Gln Met Ala Leu Glu Asp Gln Ala Ala Thr Leu Glu Tyr Lys Thr Ile
785 790 795 800

Lys Glu His Phe Ser Ser Lys Ser Pro Asn His Gly Val Asn Leu Val
805 810 815

Glu Asn Leu Asp Ser Leu Pro Pro Lys Val Pro Gln Arg Glu Ala Ser
820 825 830

Leu Gly Pro Pro Gly Ala Ser Leu Phe Gln Thr Gly Leu Ser Lys Arg
835 840 845

Leu Glu Met His His Ser Phe Ser Tyr Gly Val Asp Tyr Lys Arg Ser
850 855 860

Tyr Pro Thr Asn Ser Leu Thr Arg Ser His Gln Ala Thr Thr Leu Lys
865 870 875 880

Arg Asn Asn Thr Asn Ser Ser Asn Ser Ser His Leu Ser Arg Asn Gln
885 890 895

Ser Phe Gly Arg Gly Asp Asn Pro Pro Pro Ala Pro Gln Arg Val Asp
900 905 910

Ser Ile Gln Val His Ser Ser Gln Pro Ser Gly Gln Ala Val Thr Val
915 920 925

Ser Arg Gln Pro Ser Leu Asn Ala Tyr Asn Ser Leu Thr Arg Ser Gly
930 935 940

Leu Lys Arg Thr Pro Ser Leu Lys Pro Asp Val Pro Pro Lys Pro Ser
945 950 955 960

Phe Ala Pro Leu Ser Thr Ser Met Lys Pro Asn Asp Ala Cys Thr
965 970 975

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGCAGACTT TGGGGTCACC GGCCAGCCAC ACAGGCACCG TTTTCAGATG TCCACTTCTC 60

ATTGGGTACA TCAATCTTTT AACTTTGGGG GTCACAGTTT TAGCCACCTT TCGGGGGGTG 120

ACTGGAGCAG TAGGAGGTGT GGGGTCATTT TATGAATATA ATAAAATGGA GCTGACTATG 180

GACRRRGACT WAGTGTGGGG GAGAGGGGAC GATACAGGGT GTGTGTCTGG GAGTGCCTGG 240

GGGACAGGGA CCCCCCGGTG GTCCTATGGC AGGATGAGAA RGGAGGGACT TGGCTCCCCC 300

AGAGCCCGGT GGAAGCTACT GTTCTCTCCA GTGTCTCGAG CGTAGCCAAA ATAAGGTTGG 360

GAGGCTCCCG GCCTGTCTGC TGTGGTCTGA GCTGGCTGCA AGCCCAGGTG GGGGAGCGAG 420

TCTGGGAAGA TTGGCTTTGA CTCTCTGTTG CCAGAGGAGA TGCCATCCCA GCACGGCCCC 480

CACTGTAGTC CAGGCTCGTG GTGGCAGCGG GGGCAAGGGG AGGGGCAAGG CTGCCCCCAC 540

CCCACGCACC AAGTCACGCC AAGTCTCAGC AGGTAAAAGC ACGTGAGCCT AGGGCGAGCG 600

GAGGGAGTCC TGGTGGCCCC GCAGGTCAGG AGGGAAAGCA GGGCTCAGAG GGCATCGTGG 660

CCCCAGGGCA GGGTCCTACC TGGGGGTCAG GAGCACCTTG GTCTTGATGA TTGATTGATT 720

GATAGAATGG AGCTGGGTCT GAGCCTCCCA GGCTTGAGCT CCTGGGAGTT CTTGTGCGGT 780

GAGCTGGGCA GCTCCTGGGT AGGTCCGGGC ACCAAGCAGG CCCTGATGTG GACAGAGTCC 840

CATCAGAGGG AGCTGATGAA GAATGGTCCC TGTAAGTAAG TCACTAGGTT CAACAACTGC 900

CTGGCCGAGC ACTCAGCCCG TGGAGCTCAG GCCAACACCA GAGCCCCGGT TTTAGGGGCC 960

AGGAGAGCAG GTGACCAATT ATTTGGGGAG TCTTGGGTAG AATTTCCGCC ACACATTCTC 1020

CCCAGGGCTG CAGGGGTCTT CCGAGGCAGG GCGGTGGAGC AGGATTCAGG ATGTGGTGGG 1080

AATAGAGTGA GGGGCAGTGG GTGGGCAGAC CTGGGCGTCA GAGGTCCTGA TGGGAAAGGA 1140

GGCAGGGGCT ACCCAGAGAG GGGGGCTCGT GTGGCACAGC CCCCACCGAC TCCGCCGTCC 1200

CCCTCCCCTG TGAGCCCCGG GGGCTGTACA TACTCTACTC CATCCCCTTG TCCATCCCTG 1260

AGACCACCCC CGCCGCCCTT GCGTCGACTT AGCAACCACC TCATAGGCCC ACCCACCTCG 1320

GGATCCGAGC CAACCATCCC ACATCACAAA CTTTGGTTTG GGGGACTTTA CGTTCGTTTA 1380

ATTTCTCATT TTGTACGGAG AAATATTCTT TTCAAAAGCG TCTTTTGACT GAAGTAACTT 1440

TCCTGGTGCT GTTGTTAACT CGTTCCTTTT TTTAATTTAT TCCCCCACCC CAGGCAGCCC 1500

TCCTGGTTCC TACTCACCCT CCCCCCCTCC CCCACCCTCC GTCCCATCTG AACCATTTGT 1560

TTCTTTTCTT TCCGTCAGAT TTTGGAAAAA TTCTCCTCTC CTCCCCGCCC CCTCCACACC 1620

ATCCTCCCSG ATTTAAATAT AGTCACTGCT ACAAGTAACA GATGCACTGT GAAGATTCCA 1680

GTATTAATAA AGGTGTACTG TAATTAACAA AAAAAAAAAA AAA 1723

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 101 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Val Phe Arg Cys Pro Leu Leu Ile Gly Tyr Ile Asn Leu Leu Thr Leu
1 5 10 15

Gly Val Thr Val Leu Ala Thr Phe Arg Gly Val Thr Gly Ala Val Gly
20 25 30

Gly Val Gly Ser Phe Tyr Glu Tyr Asn Lys Met Glu Leu Thr Met Asp
35 40 45

Xaa Asp Xaa Val Trp Gly Arg Gly Asp Asp Thr Gly Cys Val Ser Gly
50 55 60

Ser Ala Trp Gly Thr Gly Thr Pro Arg Trp Ser Tyr Gly Arg Met Arg
65 70 75 80

Xaa Glu Gly Leu Gly Ser Pro Arg Ala Arg Trp Lys Leu Leu Phe Ser
85 90 95

Pro Val Ser Arg Ala
100

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 469 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCAACATACA AGCCGGCCAT ATTAGAGAGA TGGAAATAAA GCTTCCTTAA TGTTGTATAT 60

GTCTTTGAAG TACATCCGTG CATTTTTTTT TAGCATCCAA CCATTCCTCC CTTGTAGTTC 120

TCGCCCCCTC AAATCACCCT CTCCCGTAGC CCACCCGACT AACATCTCAG TCTCTGAAAA 180

TGCACAGAGA TGCCTGGCTA CCTCGCCCTG CCTTCAGCCT CACGGGGCTC AGTCTCTTTT 240

TCTCTTTGGT GCCACCAGGA CGGAGCATGG AGGTCACAGT ACCTGCCACC CTCAACGTCC 300

TCAATGGCTC TGACGCCCGC CTGCCCTGCA CCTTCAACTC CTGCTACACA GTGAACCACA 360

AACAGTTCTC CCTGAACTGG ACTTACCAGG AGTGCAACAA CTGCTCTGAG GAGATGTTCC 420

TCCAGTTCCG CATGAAGATC ATTAACCTGA AGCTGGAGCG GTTTCAAGA 469

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met His Arg Asp Ala Trp Leu Pro Arg Pro Ala Phe Ser Leu Thr Gly
1 5 10 15

Leu Ser Leu Phe Phe Ser Leu Val Pro Pro Gly Arg Ser Met Glu Val
20 25 30

Thr Val Pro Ala Thr Leu Asn Val Leu Asn Gly Ser Asp Ala Arg Leu
35 40 45

Pro Cys Thr Phe Asn Ser Cys Tyr Thr Val Asn His Lys Gln Phe Ser
50 55 60

Leu Asn Trp Thr Tyr Gln Glu Cys Asn Asn Cys Ser Glu Glu Met Phe
65 70 75 80

Leu Gln Phe Arg Met Lys Ile Ile Asn Leu Lys Leu Glu Arg Phe Gln
85 90 95

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 454 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TGGCTTTTGG CTACAGAGAG GGAAGGGAAA GCCTGAGGCC GGCATAAGGG GAGGCCTTGG 60

AACCTGAGCT GCCAATGCCA GCCCTGTCCC ATCTGCGGCC ACGATACTCG CTCCTCTCCC 120

AACAACTCCT TTGGTGGGGA CAAAAGTGAC AATTGTAGGC CAGGCACAGT GGCTCACGCC 180

TGTAATCCCA GCACTTTGGG AGGCCAAGGC GGGTGGATTA CCTCCATCTG TTTAGTAGAA 240

ATGGGCAAAA CCCCATTTTT ACTAAAAATA CAAGAATTAG CTGGGCGTGG TGGCGTGTGC 300

CTGTAATCCC AGCTATTTGG GAGGCTGAGG CAGGAGAATC GCTTGAGCCC GGGAAGCAGA 360

GGTTGCAGTG AACTGAGATA GTGATAGTGC CACTGCAATT CAGCCTGGGT GACATAGAGA 420

GACTCCATCT CAAAAAAAAA AAAAAAAAAA AAAA 454

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 736 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GTTTTTAAAC ATTATGTTCT ACATGATAAA TACATATAAT AGTATGTCTA TTTAAATAAT 60

TAAATTTGAA AAAAACTAAT CAAATATTAT CATAAGTAAT GATAAAAACC ACAATTTCTT 120

TTGCAGCAAA CTAATAACAC CTGGATTTCT CAATTTATTA AGTTGTACTT ACCTGATGCT 180

GATGATGATT ACTGTATTTA CACATTGTCT CAGAGCTCAC TCTTGCGGAG GTTGTGGCCT 240

CGAAAATGCC TTGTTGTCCC TCTGGAATCT GTCTTTTCAG CTTCATCTCC TCCTCCTCAC 300

CTCCTGCTGT GGTGCACAGA TACCTATAGG CAGGCTCCAT CTCCTCCTCC CCAGCTCCTC 360

CCCTAGTGCA CAGATACCTA TAGGCAGGCT TCATCTCCTC CTCCCCAGCT TCTCCCCTAG 420

TGCACAGATA CCTATAGGCA GGCTCCATCT CCTCCTCCCC AGCTCCTCCC CTARTGCACA 480

GACACCTATA GGCAAGCTCC ATCTCCTCCT CTTTAGCTAG CCTCCCCATC TCATCACAAC 540

GCATGTCTGT GACCTTTGGT AATCATTTAC AGTGCCACAC GGAACCCTGT ATTTTGCACA 600

CAGCAAAACA AACAATGTTT AGCTTTATTT ATGGTATTTG ATGACTGTAA ATGGAAATAA 660

ATATTGTTCT TTATTTTTTN AAAAAAAAAA AAAAAAAAAA AAANAAAAAA AAAAAAAAAA 720

AAAAAAAAAA AAAAAA 736

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 114 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Met Leu Met Met Ile Thr Val Phe Thr His Cys Leu Arg Ala His Ser
1 5 10 15

Cys Gly Gly Cys Gly Leu Glu Asn Ala Leu Leu Ser Leu Trp Asn Leu
20 25 30

Ser Phe Gln Leu His Leu Leu Leu Leu Thr Ser Cys Cys Gly Ala Gln
35 40 45

Ile Pro Ile Gly Arg Leu His Leu Leu Leu Pro Ser Ser Ser Pro Ser
50 55 60

Ala Gln Ile Pro Ile Gly Arg Leu His Leu Leu Leu Pro Ser Phe Ser
65 70 75 80

Pro Ser Ala Gln Ile Pro Ile Gly Arg Leu His Leu Leu Leu Pro Ser

85 90 95

Ser Ser Pro Xaa Ala Gln Thr Pro Ile Gly Lys Leu His Leu Leu Leu
100 105 110

Phe Ser

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1427 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTAGTTACTA ACTCCAACAC CTAATAGCAT TGGTAGAAAG CTTATAAATG CAGTTATTTA 60

GCCTCGACTA AGATTTTTCT GATACCTAGT TTCACTTTTT AATGCCCTCT GAAAGTTTTT 120

TGATCAGTTG TTTAATGGGA GATCTGAAAT GTTAAACTCA GACCAGAAAG AAGAGAACCT 180

GTTTTCTAGA AATTAGGTTT TTAATCCAAG TAAGATGCAA GCTTTTGCTT TTTTAATAAC 240

TTGTATAGCT AAAAACTTGA CGGTGAAAAG CTCTCAGATC AAAGCTGATC CTTCTGTCAG 300

TAATGATTCT AAAAATAAGC AAGATTTTAA TGGGGAATAT ATTTTATTTC ATTCTTATCT 360

CAAACCTAGG TACTGTGGTC GTTTTGAGTT CATTTCGAGG CATTTTCAAT GTGCCTCAGG 420

CCACATCCAA CCTCTYCCCA GGGCCAGATT TAATGTTCAG CCTCATAAAG GTTATCATAG 480

TTTTAACATT TAAGTACTAT TTTGCAGTGG GTATATACCA AAATTTGCTA ATAGTAAGAT 540

AACCTTAGTT ATATATCATT CACGTTAGTT CTATCTTGGA GGCAATAAAC ATTTCTTGTT 600

CAAGAAATTC ATGTTCTATC TTGGAGGCAA TAAACAAACA TTTTTTGTTC AAAATTAGGG 660

CTACCCTATT GTCCTTATGT CTTTTCCTGA TCTGTGGTCA AACATTTTTC TTAGTCATTT 720

AGAAATTTTC TATGTTGTTT TAAATTTTCT TTAAATCTAG AATGGAGTAT GTGACCAATA 780

CTTTCCTTTG GAATGGTATG GACATTTGAA ATAGAGCCCA TTCTTTATAA AGTATAAAAT 840

ATGTTTAATG CTAGTATTTT TAACTAAACT TTTGAGAAAC TAGATTCACA TGCTGTTGTA 900

AGAAATAATA CAGAGACCTC TTTCGTGTAC CTTTCACTTT GTTTCCCACA CAGTGAACAT 960

CTTTCAAAAC TGTCATACAA TATCATACCC AGGATACTGA CACTGGTATA GCTAAGATAG 1020

AGAACGTTTC CACACAGAAC TTTTTCTAGC ACAGGGATCC CTCATCTTGC TTTTGATGAC 1080

CATACCCACT TCACTCCCAT CCCTACTCCC TTCTTAACCC TTGGCAACCA TAATCTGTTC 1140

TCCATTTTTA TAGTTTTTTT TTTTTCATTT CAATAAAGCT GTATAACTGG AATCATAATA 1200

ATATGTAACC TTTTGGGATT GGCTTTTTTT CATTTAGCAT GATTTTCTGG AGGTTAATCC 1260

AGCTTATTAT GTGTATCAAG TCTATTGACA GGTACTTTTT AGTGTGAATA GAATCCCATA 1320

GTATAGATGT ACCACAGTTT GTTTAACTGT TCACCTGCTG AGAGACATTG GGCCAGTTTT 1380

TGGCTACTAT AAATAAAGTT GCTATAAACA AAAAAAAAAA AAAAAAA 1427

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 79 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Ile Leu Lys Ile Ser Lys Ile Leu Met Gly Asn Ile Phe Tyr Phe
1 5 10 15

Ile Leu Ile Ser Asn Leu Gly Thr Val Val Val Leu Ser Ser Phe Arg
20 25 30

Gly Ile Phe Asn Val Pro Gln Ala Thr Ser Asn Leu Xaa Pro Gly Pro
35 40 45

Asp Leu Met Phe Ser Leu Ile Lys Val Ile Ile Val Leu Thr Phe Lys
50 55 60

Tyr Tyr Phe Ala Val Gly Ile Tyr Gln Asn Leu Leu Ile Val Arg
65 70 75

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 572 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGCAGATTCT GTGGTTATAC TCACTCCTCA TCCCAAAGAA TGAAATTTAC CACTCTCCTC 60

TTCTTGGCAG CTGTAGCAGG GGCCCTGGTC TATGCTGAAG ATGCCTCCTC TGACTCGACG 120

GGTGCTGATC CTGCCCAGGA AGCTGGGACC TCTAAGCCTA ATGAAGAGAT CTCAGGTCCA 180

GCAGAACCAG CTTCACCCCC AGAGACAACC ACAACAGCCC AGGAGACTTC GGCGGCAGCA 240

GTTCAGGGGA CAGCCAAGGT CACCTCAAGC AGGCAGGAAC TAAACCCCCT GAAATCCATA 300

GTGGAGAAAA GTATCTTACT AACAGAACAA GCCCTTGCAA AAGCAGGAAA AGGAATGCAC 360

GGAGGCGTGC CAGGTGGAAA ACAATTCATC GAAAATGGAA GTGAATTTGC ACAAAAATTA 420

CTGAAGAAAT TCAGTCTATT AAAACCATGG GCATGAGAAG CTGAAAAGAA TGGGATCATT 480

GGACTTAAAG CCTTAAATAC CCTTGTAGCC CAGAGYTATT AAAACGAAAG CATCCAAAAA 540

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA 572

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 138 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Lys Phe Thr Thr Leu Leu Phe Leu Ala Ala Val Ala Gly Ala Leu
1 5 10 15

Val Tyr Ala Glu Asp Ala Ser Ser Asp Ser Thr Gly Ala Asp Pro Ala
20 25 30

Gln Glu Ala Gly Thr Ser Lys Pro Asn Glu Glu Ile Ser Gly Pro Ala
35 40 45

Glu Pro Ala Ser Pro Pro Glu Thr Thr Thr Thr Ala Gln Glu Thr Ser
50 55 60

Ala Ala Ala Val Gln Gly Thr Ala Lys Val Thr Ser Ser Arg Gln Glu
65 70 75 80

Leu Asn Pro Leu Lys Ser Ile Val Glu Lys Ser Ile Leu Leu Thr Glu
85 90 95

Gln Ala Leu Ala Lys Ala Gly Lys Gly Met His Gly Gly Val Pro Gly
100 105 110

Gly Lys Gln Phe Ile Glu Asn Gly Ser Glu Phe Ala Gln Lys Leu Leu
115 120 125

Lys Lys Phe Ser Leu Leu Lys Pro Trp Ala
130 135

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1223 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCTTGTTCCA CGTAGCTGGC AAGGTCTTCA TTCACTTGCC ACTGCTAGTC TTCCAACCCT 60

TCTGGACTTT CTTTGCTCTT GTCTTGTTTT GGGTGTACTG GATCATGACA CTTCTTTTTC 120

TTGGCACTAC CGGCAGTCCT GTTCAGAATG AGCAAGGCTT TGTGGAGTTC AAAATTTCTG 180

GGCCTCTGCA GTACATGTGG TGGTACCATG TGGTGGGCCT GATTTGGATC AGTGAATTTA 240

TTCTAGCATG TCAGCAGATG ACAGTGGCAG GAGCTGTGGT AACATACTAT TTTACTAGGG 300

ATAAAAGGAA TTTGCCATTT ACACCTATTT TGGCATCAGT AAATCGCCTT ATYCGTTACC 360

ACCTAGGTAC GGTGGCAAAA GGATCTTTCA TTATCACATT AGTCAAAATT CCGCGAATGA 420

TCCTTATGTA TATTCACAGT CAGCTCAAAG GAAAGGAAAA TGCTTGTGCA CGATGTGTGC 480

TGAAATCTTG CATTTGTTGC CTTTGGTGTC TTGAAAAGTG CCTAAATTAT TTAAATCAGA 540

ATGCATACAC AGCCACAGCT ATCAACAGCA CCAACTTCTG CACCTCAGCA AAGGATGCCT 600

TTGTCATTCT GGTGGAGAAT GCTTTGCGAG TGGCTACCAT CAACACAGTA GGAGATTTTA 660

TGTTATTCCT TGGCAAGGTG CTGATAGTCT GCAGCACAGG TTTAGCTGGG ATTATGCTGC 720

TCAACTACCA GCAGGACTAC ACAGTATGGG TGCTGCCTCT GATCATCGTC TGCCTCTTTG 780

CTTTCCTAGT CGCTCATTGC TTCCTGTCTA TTTATGAAAT GGTAGTGGAT GTATTATTCT 840

KGKGTTTTGC CATTGAWACA AAATACAATG ATGGGMGCCC TGGCAGAGAA TTCTATATGG 900

ATAAAGTGCT GATGGAGTTT GTGGAAAACA GTAGGAAAGC AATGAAAGAA GCTGGTAAGG 960

GAGGCGTCGC TGATTCCAGA GAGCTAAAGC CGATGCTGAA GAAAAGGTGA CTGGTCTCAT 1020

GAGCCCTGAA GAATGAACTC AGAGGAGGTT GTTTACATGA GGTTCTCCCA CTCACCAGCT 1080

GTTGAGAGTC TGCGATTATG AAGAGCAGGA TCTTATTACT TCAATGAAAG CATGTAACAA 1140

GTTTCTCAAA CCACCAACAG CCAAGTGGAT TTGGTACAGT GCGGCTGTCT AATAAATAAT 1200

CAAAAGCAAA AAAAAAAAAA AAA 1223

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 301 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Thr Leu Leu Phe Leu Gly Thr Thr Gly Ser Pro Val Gln Asn Glu
1 5 10 15

Gln Gly Phe Val Glu Phe Lys Ile Ser Gly Pro Leu Gln Tyr Met Trp
20 25 30

Trp Tyr His Val Val Gly Leu Ile Trp Ile Ser Glu Phe Ile Leu Ala
35 40 45

Cys Gln Gln Met Thr Val Ala Gly Ala Val Val Thr Tyr Tyr Phe Thr
50 55 60

Arg Asp Lys Arg Asn Leu Pro Phe Thr Pro Ile Leu Ala Ser Val Asn
65 70 75 80

Arg Leu Ile Arg Tyr His Leu Gly Thr Val Ala Lys Gly Ser Phe Ile
85 90 95

Ile Thr Leu Val Lys Ile Pro Arg Met Ile Leu Met Tyr Ile His Ser
100 105 110

Gln Leu Lys Gly Lys Glu Asn Ala Cys Ala Arg Cys Val Leu Lys Ser
115 120 125

Cys Ile Cys Cys Leu Trp Cys Leu Glu Lys Cys Leu Asn Tyr Leu Asn
130 135 140

Gln Asn Ala Tyr Thr Ala Thr Ala Ile Asn Ser Thr Asn Phe Cys Thr
145 150 155 160

Ser Ala Lys Asp Ala Phe Val Ile Leu Val Glu Asn Ala Leu Arg Val
165 170 175

Ala Thr Ile Asn Thr Val Gly Asp Phe Met Leu Phe Leu Gly Lys Val
180 185 190

Leu Ile Val Cys Ser Thr Gly Leu Ala Gly Ile Met Leu Leu Asn Tyr
195 200 205

Gln Gln Asp Tyr Thr Val Trp Val Leu Pro Leu Ile Ile Val Cys Leu
210 215 220

Phe Ala Phe Leu Val Ala His Cys Phe Leu Ser Ile Tyr Glu Met Val
225 230 235 240

Val Asp Val Leu Phe Xaa Xaa Phe Ala Ile Xaa Thr Lys Tyr Asn Asp
245 250 255

Gly Xaa Pro Gly Arg Glu Phe Tyr Met Asp Lys Val Leu Met Glu Phe
260 265 270

Val Glu Asn Ser Arg Lys Ala Met Lys Glu Ala Gly Lys Gly Gly Val
275 280 285

Ala Asp Ser Arg Glu Leu Lys Pro Met Leu Lys Lys Arg
290 295 300

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2460 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GAGAAGCGCG ATGACGGCTA CGAGGCCGCT GCCAGCTCCA AAACTAGCTC AGGCGATGCC 60

TCCTCACTCA GCATCGAGGG AGACTAACAA ACTCCGGGCA AAGTTGGGGC TGAAACCCTT 120

GGAGGTTAAT GCCATCAAGA AGGAGGCGGG CACCAAGGAG GAGCCCGTGA CAGCTGATGT 180

CATCAACCCT ATGGCCTTGC GACAGCGAGA GGAGCTGCGG GAGAAGCTGG CGGCTGCCAA 240

GGAGAAGCGC CTGCTGAACC AAAAGCTGGG GAAGATAAAG ACCCTAGGAG AGGATGACCC 300

CTGGCTGGAC GACACTGCAG CCTGGATCGA GAGGAGCCGG CAGCTGCAGA AGGAGAAGGA 360

CCTGGCAGAG AAGAGGGCCA AGTTACTGGA GGAGATGGAC CAAAAGTTTG GTGTCAGCAC 420

TCTGGTGGAG GAGGAGTTCG GGCAGAGGCG GCAGGACCTG TACAGTGCCC GGGACCTGCA 480

GGGCCTCACT GTGGAGCATG CCATTGATTC CTTCCGAGAA GGGGAGACAA TGATTCTTAC 540

CCTCAAGGAC AAAGGCGTGC TGCAGGAGGA GGAGGACGTG CTGGTGAACG TGAACCTGGT 600

GGATAAGGAG CGGGCAGAGA AAAATGTGGA GCTGCGGAAG AAGAAGCCTG ACTACCTGCC 660

CTATGCCGAG GACGAGAGCG TGGACGACCT GGCGCAGCAA AAACCTCGCT CTATCCTGTC 720

CAAGTATGAC GAAAAGCTTG AAGGGGAGCG GCCACATTCC TTCCGCTTGG AGCAGGGCGG 780

CACGGCTGAT GGCCTGCGGG AGCGGGAGCT GGAGGAGATC CGGGCCAAGC TGCGGCTGCA 840

GGCTCAGTCC CTGAGCACAG TGGGGCCCCG GCTGGCCTCC GAATACCTCA CGCCTGAGGA 900

GATGGTGACC TTTAAAAAGA CCAAGCGGAG GGTGAAGAAA ATCCGCAAGA AGGAGAAGGA 960

GGTAGTAGTG CGGGCAGATG ACTTGCTGCC TCTCGGGGAC CAGACTCAGG ATGGGGACTT 1020

TGGTTCCAGA CTGCGGGGAC GGGGTCGCCG CCGAGTGTCC GAAGTGGAGG AGGAGAAGGA 1080

GCCTGTGCCT CAGCCCCTGC CGTCGGACGA CACCCGAGTG GAGAACATGG ACATCAGTGA 1140

TGAGGAGGAA GGTGGAGCTC CACCGCCGGG GTCCCCGCAG GTGCTGGAGG AGGACGAGGC 1200

GGAGCTGGAG CTGCAGAAGC AGCTGGAGAA GGGACGCCGG CTGCGACAGT TACAGCAGCT 1260

ACAGCAGCTG CGAGACAGTG GCGAGAAGGT GGTGGAGATT GTGAAGAAGC TGGAGTCTCG 1320

CCAGCGGGGC TGGGAGGAGG ATGAGGATCC CGAGCGGAAG GGGGCCATCG TGTTCAACGC 1380

CACGTCCGAG TTCTGCCGCA CCTTGGGGGA GATCCCCACC TACGGGCTGG CTGGCAATCC 1440

CGAGGAGCAG GAGGAGCTCA TGGACTTTGA ACGGGATGAG GAGCGCTCAG CCAACGGTGG 1500

CTCCGAATCT GACGGGGAGG AGAACATCGG CTGGAGCACG GTGAACCTGG ACGAGGAGAA 1560

GCAGCAGCAG GATGTGAGGG CCACGCCGCT GGGGGGTGGG CGTTTGGGGG TGCTCAAGCT 1620

GGAGATGAGC ACCGGGCTCG GTGTCCAGAG CCTCAGCCTC CTCATCCAGA GTGGGCTCTG 1680

CAGACCTCCC AGGGCGATCT GAGGAGTAAA TGAGGAAATT AAATGTTGTG GAGGGCTGGT 1740

GCCTGGCAGG TGGTGACCAG TGGGTGGGGC TGAGAAGAGC CGGTATGGCC TGCTAACCAC 1800

CCCCGCCACG TGTCCCGTAG TTCTCTGCTT CCTCCACCAC CATCCTGGAC GAGGAACCGA 1860

TCGTGAATAG GGGGCTGGCA GCTGCCCTGC TCCTGTGTCA GAACAAAGGG CTGCTGGAGA 1920

CCACAGTGCA GAAGGTGGCC CGGGTGAAGG CCCCCAACAA GTCGCTGCCC TCAGCCGTGT 1980

ACTGCATCGA GGATAAGATG GCCATCGATG ACAAGTACAG CCGGAGGGAG GAATACCGAG 2040

GCTTCACACA GGACTTCAAG GAGAAGGACG GTACAAACCC GACGTTAAGA TCGAATACGT 2100

GGATGAGACG GGCCGGAAAC TCACACCCAA GGAGGCTTTC CGGCAGCTGT CGCACCGCTT 2160

CCATGGCAAG GGCTCAGGCA AGATGAAGAC AGAGCGGCGG ATGAAGAAGC TGGACGAGGA 2220

GGCGCTCCTG AAGAAGATGA GCTCCAGSGA CACGCCCYTG GGCACCGTGG CCCTGYTCCA 2280

GGAGAAGCAG AAGGCTCAGA AGACCCCCTA CATYGTGTTC AGCGGCAGCG GCAAGAGCAT 2340

GAACGCGAAC ACCATCACCA AGTGACAGCG CCCTCCCGCC CCGGCCCTGC CTCAACCTTC 2400

ATATTAAATA AAGCTCCCTC CTTAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 2460

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 563 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Thr Ala Thr Arg Pro Leu Pro Ala Pro Lys Leu Ala Gln Ala Met
1 5 10 15

Pro Pro His Ser Ala Ser Arg Glu Thr Asn Lys Leu Arg Ala Lys Leu
20 25 30

Gly Leu Lys Pro Leu Glu Val Asn Ala Ile Lys Lys Glu Ala Gly Thr
35 40 45

Lys Glu Glu Pro Val Thr Ala Asp Val Ile Asn Pro Met Ala Leu Arg
50 55 60

Gln Arg Glu Glu Leu Arg Glu Lys Leu Ala Ala Ala Lys Glu Lys Arg
65 70 75 80

Leu Leu Asn Gln Lys Leu Gly Lys Ile Lys Thr Leu Gly Glu Asp Asp
85 90 95

Pro Trp Leu Asp Asp Thr Ala Ala Trp Ile Glu Arg Ser Arg Gln Leu
100 105 110

Gln Lys Glu Lys Asp Leu Ala Glu Lys Arg Ala Lys Leu Leu Glu Glu
115 120 125

Met Asp Gln Lys Phe Gly Val Ser Thr Leu Val Glu Glu Glu Phe Gly
130 135 140

Gln Arg Arg Gln Asp Leu Tyr Ser Ala Arg Asp Leu Gln Gly Leu Thr
145 150 155 160

Val Glu His Ala Ile Asp Ser Phe Arg Glu Gly Glu Thr Met Ile Leu
165 170 175

Thr Leu Lys Asp Lys Gly Val Leu Gln Glu Glu Glu Asp Val Leu Val
180 185 190

Asn Val Asn Leu Val Asp Lys Glu Arg Ala Glu Lys Asn Val Glu Leu
195 200 205

Arg Lys Lys Lys Pro Asp Tyr Leu Pro Tyr Ala Glu Asp Glu Ser Val
210 215 220

Asp Asp Leu Ala Gln Gln Lys Pro Arg Ser Ile Leu Ser Lys Tyr Asp
225 230 235 240

Glu Lys Leu Glu Gly Glu Arg Pro His Ser Phe Arg Leu Glu Gln Gly
245 250 255

Gly Thr Ala Asp Gly Leu Arg Glu Arg Glu Leu Glu Glu Ile Arg Ala
260 265 270

Lys Leu Arg Leu Gln Ala Gln Ser Leu Ser Thr Val Gly Pro Arg Leu
275 280 285

Ala Ser Glu Tyr Leu Thr Pro Glu Glu Met Val Thr Phe Lys Lys Thr
290 295 300

Lys Arg Arg Val Lys Lys Ile Arg Lys Lys Glu Lys Glu Val Val Val
305 310 315 320

Arg Ala Asp Asp Leu Leu Pro Leu Gly Asp Gln Thr Gln Asp Gly Asp
325 330 335

Phe Gly Ser Arg Leu Arg Gly Arg Gly Arg Arg Arg Val Ser Glu Val
340 345 350

Glu Glu Glu Lys Glu Pro Val Pro Gln Pro Leu Pro Ser Asp Asp Thr
355 360 365

Arg Val Glu Asn Met Asp Ile Ser Asp Glu Glu Glu Gly Gly Ala Pro
370 375 380

Pro Pro Gly Ser Pro Gln Val Leu Glu Glu Asp Glu Ala Glu Leu Glu
385 390 395 400

Leu Gln Lys Gln Leu Glu Lys Gly Arg Arg Leu Arg Gln Leu Gln Gln
405 410 415

Leu Gln Gln Leu Arg Asp Ser Gly Glu Lys Val Val Glu Ile Val Lys

420 425 430

Lys Leu Glu Ser Arg Gln Arg Gly Trp Glu Glu Asp Glu Asp Pro Glu
435 440 445

Arg Lys Gly Ala Ile Val Phe Asn Ala Thr Ser Glu Phe Cys Arg Thr
450 455 460

Leu Gly Glu Ile Pro Thr Tyr Gly Leu Ala Gly Asn Arg Glu Glu Gln
465 470 475 480

Glu Glu Leu Met Asp Phe Glu Arg Asp Glu Glu Arg Ser Ala Asn Gly
485 490 495

Gly Ser Glu Ser Asp Gly Glu Glu Asn Ile Gly Trp Ser Thr Val Asn
500 505 510

Leu Asp Glu Glu Lys Gln Gln Gln Asp Val Arg Ala Thr Pro Leu Gly
515 520 525

Gly Gly Arg Leu Gly Val Leu Lys Leu Glu Met Ser Thr Gly Leu Gly
530 535 540

Val Gln Ser Leu Ser Leu Leu Ile Gln Ser Gly Leu Cys Arg Pro Pro
545 550 555 560

Arg Ala Ile

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ANAGGCTCCTC CATTCCTACA GCCATCTT 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CNGTCCAACTG CTTGTAGGTT ATAGCAGA 29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TNCTCTACTTC ACCCTTTTCG GTGCATCG 29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TNCATAAAAT GACCCCACAC CTCCTACTG 29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GNGAGGTAGCC AGGCATCTCT GTGCATTT 29

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ANCACAGCAGG AGGTGAGGAG GAGGAGAT 29

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CNTTGATCTGA GAGCTTTTCA CCGTCAAG 29

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

ANCTGCCAAGA AGAGGAGAGT GGTAAATT 29

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ANGAAAAAGAA GTGTCATGAT CCAGTACA 29

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TNATACTTGGA CAGGATAGAG CGAGGTTT 29

What is claimed is:

1. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:2;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:2 from nucleotide 41 to nucleotide 760;

(c) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CB107_1 deposited under accession number ATCC 98279;

(d) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279;

(e) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CB107_1 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279;

(g) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:3;

(h) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:3 having biological activity;

(i) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(f) above;

(j) a polynucleotide which encodes a species homologue of the protein of (g) or (h) above ; and

(k) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h).

2. A composition of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with a composition of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by a composition of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:3;

(b) the amino acid sequence of SEQ ID NO:3 from amino acid 127 to amino acid 240;

(c) fragments of the amino acid sequence of SEQ ID NO:3; and

(d) the amino acid sequence encoded by the cDNA insert of clone CB107_1 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

9. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:3.

10. The composition of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:3 from amino acid 127 to amino acid 240.

11. The composition of claim 8, further comprising a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:2, SEQ ID NO:1 or SEQ ID NO:4 .

14. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 374 to nucleotide 1108;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 500 to nucleotide 1108;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 1 to nucleotide 387;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CG300_3 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CG300_3 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

15. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 23 to amino acid 57;

(c) fragments of the amino acid sequence of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone CG300_3 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

17. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 126 to nucleotide 3053;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 180 to nucleotide 3053;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 49 to nucleotide 382;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CJ145_1 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CJ145_1 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

18. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 87;

(c) fragments of the amino acid sequence of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone CJ145_1 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

20. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 40 to nucleotide 342;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 127 to nucleotide 342;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 11 to nucleotide 181;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CJ160_11 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CJ160_11 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CJ160_11 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CJ160_11 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

21. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 7 to amino acid 48;

(c) fragments of the amino acid sequence of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone CJ160_11 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

23. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 180 to nucleotide 467;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 267 to nucleotide 467;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO20_1 deposited under accession number ATCC 98279;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO20_1 deposited under accession number ATCC 98279;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 1 to amino acid 37;

(c) fragments of the amino acid sequence of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO20_1 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11 or SEQ ID NO:13.

26. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 176 to nucleotide 520;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 317 to nucleotide 520;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:14 from nucleotide 118 to nucleotide 413;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO223_3 deposited under accession number ATCC 98291;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO223_3 deposited under accession number ATCC 98291;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:15;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:15 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

27. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:15;

(b) the amino acid sequence of SEQ ID NO:15 from amino acid 1 to amino acid 80;

(c) fragments of the amino acid sequence of SEQ ID NO:15; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO223_3 deposited under accession number ATCC 98291;

the protein being substantially free from other mammalian proteins.

28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:14.

29. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 303 to nucleotide 542;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:16 from nucleotide 1 to nucleotide 435;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CO310_2 deposited under accession number ATCC 98279;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CO310_2 deposited under accession number ATCC 98279;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:17;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:17 having biological activity;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

30. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:17;

(b) the amino acid sequence of SEQ ID NO:17 from amino acid 1 to amino acid 44;

(c) fragments of the amino acid sequence of SEQ ID NO:17; and

(d) the amino acid sequence encoded by the cDNA insert of clone CO310_2 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:16.

32. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 40 to nucleotide 455;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 85 to nucleotide 455;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:18 from nucleotide 265 to nucleotide 515;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CP258_3 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CP258_3 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:19;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:19 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

33. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:19;

(b) the amino acid sequence of SEQ ID NO:19 from amino acid 64 to amino acid 138;

(c) fragments of the amino acid sequence of SEQ ID NO:19; and

(d) the amino acid sequence encoded by the cDNA insert of clone CP258_3 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:18.

35. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 105 to nucleotide 1007;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 801 to nucleotide 1007;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:20 from nucleotide 1 to nucleotide 352;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CW1155_3 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CW1155_3 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:21;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:21 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

36. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:21;

(b) the amino acid sequence of SEQ ID NO:21 from amino acid 1 to amino acid 83;

(c) fragments of the amino acid sequence of SEQ ID NO:21; and

(d) the amino acid sequence encoded by the cDNA insert of clone CW1155_3 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:20.

38. A composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 11 to nucleotide 1699;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 1682 to nucleotide 1699;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:22 from nucleotide 737 to nucleotide 1134;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone CZ247_2 deposited under accession number ATCC 98279;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone CZ247_2 deposited under accession number ATCC 98279;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:23;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:23 having biological activity;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

39. A composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:23;

(b) the amino acid sequence of SEQ ID NO:23 from amino acid 298 to amino acid 374;

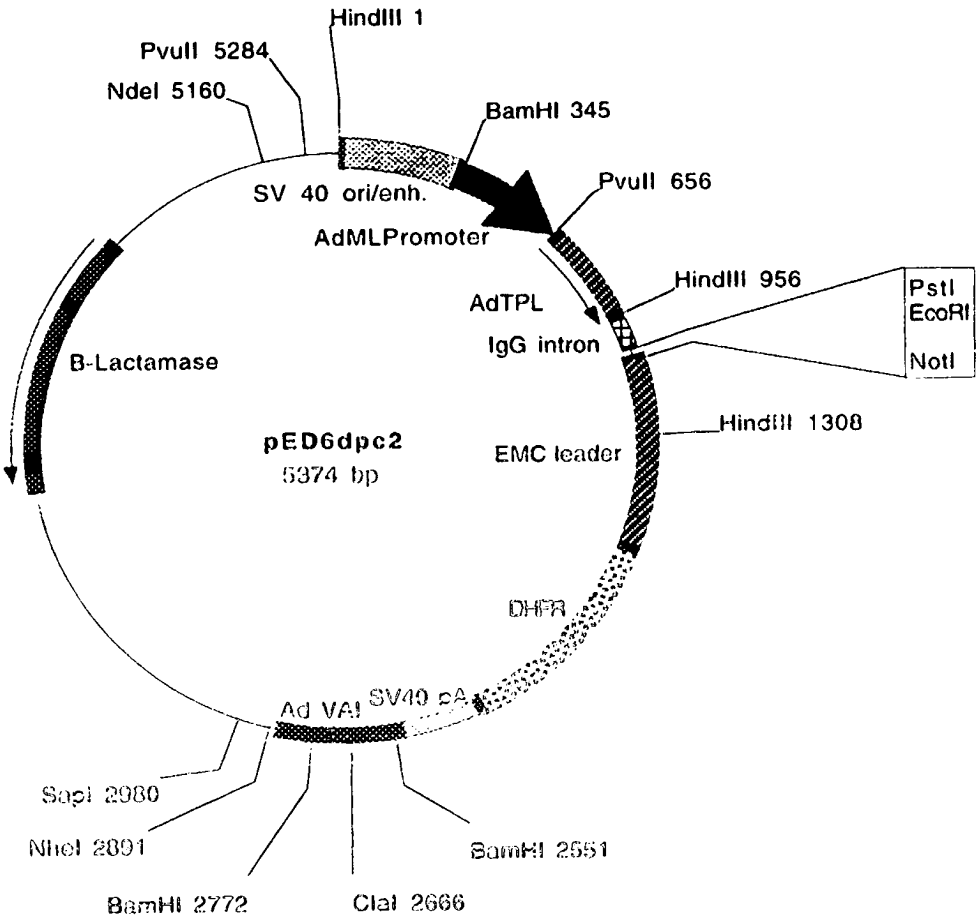
(c) fragments of the amino acid sequence of SEQ ID NO:23; and

(d) the amino acid sequence encoded by the cDNA insert of clone CZ247_2 deposited under accession number ATCC 98279;

the protein being substantially free from other mammalian proteins.

40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:22.

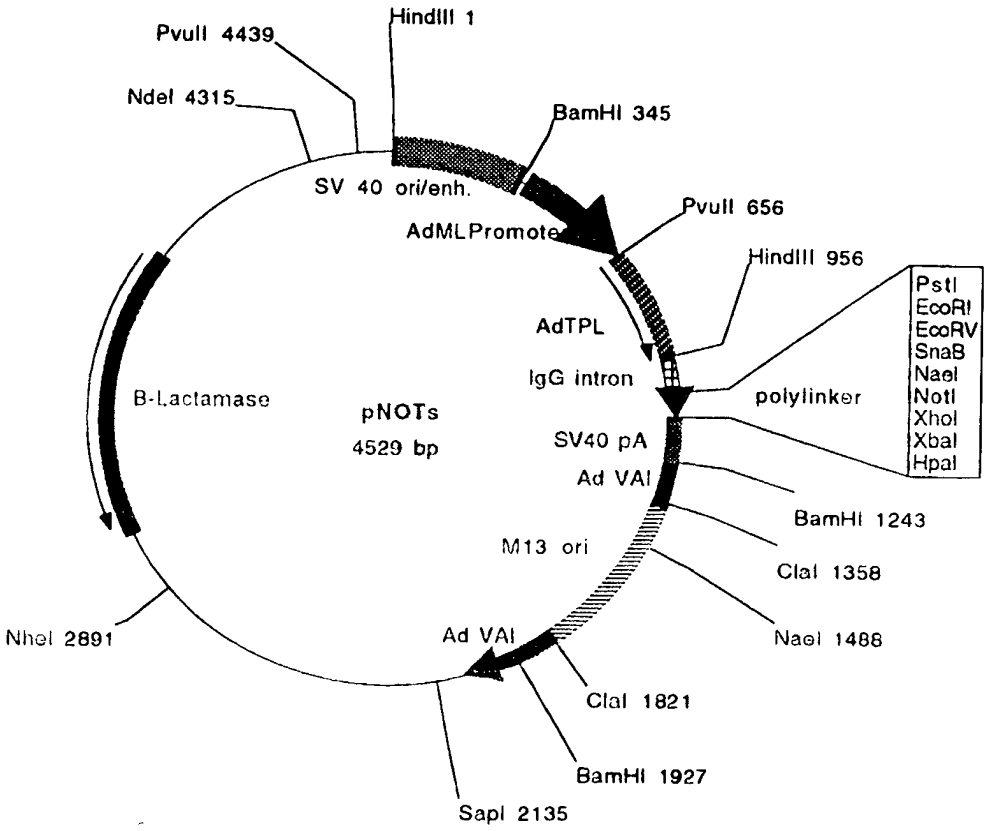
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

PCT

INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/11, C07K 14/47, A61K 38/00

A3

(11) International Publication Number: WO 98/27205

(43) International Publication Date: 25 June 1998 (25.06.98)

(21) International Application Number: PCT/US97/23330

(22) International Filing Date: 17 December 1997 (17.12.97)

(30) Priority Data:

08/769,192	18 December 1996 (18.12.96)	US
08/783,401	13 January 1997 (13.01.97)	US
08/991,872	16 December 1997 (16.12.97)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
29 October 1998 (29.10.98)

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Interna[illegible] Application No
PCT/US 97/23330

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/11 C07K14/47 A61K38/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WATANABE T K ET AL: "Cloning and characterization of two novel human cDNAs (NELL1 and NELL2) encoding proteins with six EGF-like repeats." GENOMICS 38 (3). 1996. 273-276. ISSN: 0888-7543, XP002064052 see the sequence of the Nell human gene ---	1-13
P,A	EP 0 796 913 A (OTSUKA PHARMA CO LTD) 24 September 1997 see the sequences SEQ ID NO: 37-39 -----	1-13

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 May 1998	08. 09. 98
Name and mailing address of the ISA European Patent Office, P B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Halle, F

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 97/ 23330

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See attached sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1 - 13

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-13

Polynucleotide and amino acid sequences (SEQ ID NO: 1-4) of the clone "CB107 1"

2. Claims: 14-16

Polynucleotide and amino acid sequences (SEQ ID NO: 5 and 6) of the clone "CG300 3"

3. Claims: 17-19

Polynucleotide and amino acid sequences (SEQ ID NO: 7 and 8) of the clone "CJ145 1"

4. Claims: 20-22

Polynucleotide and amino acid sequences (SEQ ID NO: 9 and 10) of the clone "CJ160 11"

5. Claims: 23-25

Polynucleotide and amino acid sequences (SEQ ID NO: 11-13) of the clone "CO20 1"

6. Claims: 26-28

Polynucleotide and amino acid sequences (SEQ ID NO: 14 and 15) of the clone "CO223 3"

7. Claims: 29-31

Polynucleotide and amino acid sequences (SEQ ID NO: 16 and 17) of the clone "CO310 2"

8. Claims: 32-34

Polynucleotide and amino acid sequences (SEQ ID NO: 18 and 19) of the clone "CP258 3"

9. Claims: 35-37

Polynucleotide and amino acid sequences (SEQ ID NO: 20 and 21) of the clone "CW1155 3"

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: 38-40

Polynucleotide and amino acid sequences (SEQ ID NO: 22 and 23) of the clone "CZ247 2"

INTERNATIONAL SEARCH REPORT

Information on patent family members

Intern[redacted] Application No
PCT/US 97/23330

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0796913 A	24-09-1997	JP 9308492 A	02-12-1997
		CA 2200371 A	19-09-1997